



## Diurnal self-aggregation

Härter, Jan Olaf Mirko; Meyer, Bettina; Nissen, Silas Boye

*Published in:*  
n p j Climate and Atmospheric Science

*DOI:*  
[10.1038/s41612-020-00132-z](https://doi.org/10.1038/s41612-020-00132-z)

*Publication date:*  
2020

*Document version*  
Publisher's PDF, also known as Version of record

*Citation for published version (APA):*  
Härter, J. O. M., Meyer, B., & Nissen, S. B. (2020). Diurnal self-aggregation. *n p j Climate and Atmospheric Science*, 3, [30]. <https://doi.org/10.1038/s41612-020-00132-z>

## ARTICLE OPEN



## Diurnal self-aggregation

Jan O. Haerter<sup>1,2,3✉</sup>, Bettina Meyer<sup>1</sup> and Silas Boye Nissen<sup>1</sup>

Convective self-aggregation is a modelling paradigm for convective rain cell organisation over a constant-temperature tropical sea surface. This set-up can give rise to cloud clusters developing over timescales of weeks. In reality, sea-surface temperatures do oscillate diurnally, affecting the atmospheric state and influencing rain rates significantly. Over land, surface temperatures vary more strongly. Here, we carry out a suite of cloud-resolving numerical experiments, and find that qualitatively different dynamics emerge from modest surface temperature oscillations: while the spatial distribution of rainfall is homogeneous during the first day, already on the second day, the rain field is firmly structured. In later days, this clustering becomes stronger and alternates from day to day. We show that these features are robust to changes in resolution, domain size and mean surface temperature, but can be removed by a reduction of the amplitude of diurnal surface temperature oscillation, suggesting a transition from a random to a clustered state. Maximal clustering occurs at a scale of  $l_{\text{max}} \approx 180$  km, which we relate to the emergence of mesoscale convective systems. At  $l_{\text{max}}$ , rainfall is strongly enhanced and far exceeds the rainfall expected at random. Simple conceptual modelling helps interpret the transition to clustering, which is driven by the formation of mesoscale convective systems, and brings about day-to-day moisture oscillations. Our results may help clarify how continental extremes build up, and how cloud clustering over the tropical ocean could emerge as an instance of spontaneous symmetry breaking at timescales much faster than in conventional radiative–convective equilibrium self-aggregation.

*npj Climate and Atmospheric Science* (2020)3:30; <https://doi.org/10.1038/s41612-020-00132-z>

## INTRODUCTION

Due to their relatively low resolution, current general circulation models cannot simulate mesoscale convective organisation explicitly<sup>1</sup>. Yet, the recent observed positive trend in regional tropical rainfall was stated to predominantly stem from changes in the frequency of organised deep convection<sup>2</sup>. Similarly, in mid-latitudes, the majority of flood-producing rain was attributed to mesoscale convective systems (MCSs)<sup>3,4</sup>. In line with Moeng and LeMone<sup>5</sup>, we here take MCSs to be long-lived complexes of rain cells spanning  $\sim 100$  km in diameter. Over continental regions, the clustering of convection in an MCS poses a risk to humans as severe storms can lead to intense downdraughts and flash flooding<sup>3,4,6,7</sup>. Pronounced long-lived clustering can be simulated in a process known as convective self-aggregation forming over the timescale of weeks<sup>8–10</sup>. There, the radiative–convective equilibrium (RCE) experimental set-up<sup>11,12</sup> is usually employed, assuming spatially and temporally constant surface temperature ( $\sim 300$  K). In RCE self-aggregation, radiation feedbacks have emerged as the “smoking gun” for sustaining and increasing clustering<sup>8</sup>. Still, factors such as sea-surface temperature feedbacks<sup>13</sup>, domain size, geometry and resolution<sup>10</sup>, as well as periodically varying insolation<sup>14,15</sup> or cold-pool effects<sup>16,17</sup> were also stated to influence RCE self-aggregation.

Prescribing constant boundary conditions is an elegant model simplification, but it is not always realistic. Diurnal sea-surface temperature amplitudes can be as large as two<sup>18,19</sup> to five<sup>20</sup> Kelvin, especially under weak surface wind conditions and strong insolation. Diurnal surface temperature amplitudes of this magnitude change the properties of the atmosphere<sup>20</sup>, resulting in diurnally varying precipitation intensity and affecting large-scale phenomena, such as the Madden–Julian Oscillation (MJO) or the El Niño phenomenon. Satellite observations show a diurnal cycle in cloud height for the MJO<sup>21,22</sup>. Diurnally varying

precipitation intensity also emerges from numerical simulations that account for differences in insolation<sup>23</sup>. Further, in contrast to RCE self-aggregation simulations, observed tropical cloud clusters typically are more transient, as they often persist for less than two days<sup>24</sup>.

Mechanistically, cold pools (CPs), that is, denser air formed by rain evaporation under convective clouds, were long implicated in the organisation of convection<sup>25,26</sup>, both by thermodynamic and mechanical effects<sup>27–31</sup>, and were suggested to lead to clustering<sup>32,33</sup>. In the following, when referring to MCS, we take this to encompass both the cluster of deep convective cells and the associated CPs, which are often long-lived. The continental diurnal cycle of deep convection is characterised by an afternoon or evening peak in rainfall<sup>31,34–38</sup>. Over the ocean, this daily period can further be modulated by bi-diurnal oscillations, a dynamics observed by Chen and Houze for local cloudiness in MCSs and referred to as “diurnal dancing”<sup>39</sup>.

Previous studies of convective self-organisation under diurnally varying surface temperature either did not address the emergence of MCSs or simulated a relatively small domain<sup>37</sup>. Current state-of-the-art regional climate models are now capable of simulating much larger areas (beyond 1000 km horizontally at kilometre resolution)<sup>40–43</sup>. Encompassing features, such as terrain variation and inhomogeneous external forcing, adds unique regional insight but makes mechanistic analysis of self-organised clustering more cumbersome.

Through an idealised set-up, we analyse the multi-day organisation in the convective rain field over land and sea. In our set-up, surface heating is spatially homogeneous but oscillates diurnally. Our focus is on the spontaneous emergence of clustering. We show that a transition from a homogeneous to a strongly clustered state occurs spontaneously when the amplitude of the diurnal surface temperature oscillation is sufficiently large.

<sup>1</sup>Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark. <sup>2</sup>Complexity and Climate, Leibniz Centre for Tropical Marine Research, Fahrenheitstrasse 6, 28359 Bremen, Germany. <sup>3</sup>Physics and Earth Sciences, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany. ✉email: haerter@nbi.ku.dk

We refer to the emergent convective clusters as MCSs because of their typical scale of  $\sim 100$  km<sup>44</sup> and term the clustering process diurnal self-aggregation due to the lack of imposed spatial scales. This clustering oscillates bi-diurnally and continues to strengthen from day to day, features we back by a conceptual model. Finally, we discuss the relevance for the emergence of extreme continental rainfall and the implications for organised convection over the tropical ocean.

## RESULTS

We carry out a suite of numerical experiments on square mesoscale model domains ( $L \times L$ ) of up to  $L = 960$  km horizontal length and horizontal grid resolutions of 1 km and finer. Diurnally oscillating, spatially homogeneous surface temperature,  $T_s(t)$ , is prescribed as

$$T_s(t) = \overline{T_s} - T_a \cos(2\pi t/t_0), \quad (1)$$

where  $\overline{T_s} = 298$  K,  $t_0 \equiv 24$  h is the duration of the simulated model day, and  $\overline{T_s}$  and  $T_a$  represent the temporal average and amplitude of  $T_s(t)$ , respectively. The diurnal insolation cycle is chosen to be typical for the equator and peaks at noon (just like the surface temperature  $T_s$ ). We provide sensitivity experiments exploring resolution, domain size, rain evaporation, surface conditions and insolation (see details in “Methods”). We contrast simulations with different diurnal surface temperature amplitudes,  $T_a$ , but equal average surface temperature, and refer to experiments with  $T_a = 2$  K, 3.5 K and 5 K as A2, A3.5 and A5, respectively. Modifiers, such as A5b or A5sea (see Table 1), label sensitivity studies. Unless explicitly stated, our key results regarding clustering occurring under varying values of  $T_a$  are qualitatively robust under the sensitivity experiments. Each numerical simulation is run for several days, allowing for a spin-up and quasi-steady-state period (Table 1; Supplementary Fig. 1).

### Domain-mean time series

Unsurprisingly, the differences in diurnal surface temperature amplitudes are reflected in larger diurnal amplitudes of atmospheric near-surface temperatures (Fig. 1a; Supplementary Fig. 1a, b). Changes in the time series of domain-mean rain rate are more profound (Fig. 1b): whereas A5 yields a relatively sharp mid-afternoon single-peak structure, the curve transitions to a broader and double-peaked structure for A2, where it approximates the diurnal cycle typical of oceanic convection<sup>45</sup>. For  $T_a = 0$ , only a weak nocturnal maximum remains (Supplementary Fig. 2f, compare: Janowiak et al.)<sup>46</sup>. Again, the differences in the temporal mean (horizontal lines) are minimal, reflecting radiation constraints on rainfall<sup>47</sup>. Real surface temperatures normally peak at a delay relative to local solar noon. Such delays accordingly shift the onset of organisation; hence, radiative effects are not key in this timing (Supplementary Fig. 2e). Curves that vary in a similar manner to the rain rate are found for rain area fraction, which, by contrast, differs from those of rain rate immediately before the central peak (Fig. 1c). These differences are made more transparent when inspecting rain rates conditional on a threshold (Fig. 1d). Both mean and heavy precipitation show a pronounced evening peak for A5 and an early-morning or nocturnal peak for A2. In summary, whereas time averages of rain rate are nearly identical for numerical experiments with varying diurnal surface temperature amplitudes, the time series differ markedly.

### Quantifying clustering

Do the differences in domain-mean time series hint at differences in spatial organisation? Consider the horizontal patterns formed by surface rainfall, visualised by temporally averaging the rainfall pattern during each model day (Fig. 1e, f). During the spin-up from the initial condition (day 1), both A2 and A5 show modest and

**Table 1.** Summary of numerical experiments.

Experiment Name	Diurnal surface temperature Amplitude, $T_a$ [K]	Horizontal resolution $dx$ [km]	Domain size $L$ [km]	Days with 3D output
A5a	5	1	960	1–6
A2a	2	1	960	1, 4–6, 8
A5b	5	1	480	1–8
A2b	2	1	480	1–4, 8*, 9*
A5c	5	0.5	240	1–3
A5d	5	0.2	240	1–3
A3.5	3.5	1	480	3–5
A5sea	5	1	480	1–4
A2sea	2	1	480	1, 8
A5vent05	5	1	480	4
A5vent001	5	1	480	4
A5p2K	5	1	480	4, 8
A5ph3.6h	5	1.5	384	3–5
A5ph6h	5	1.5	384	3–5
A0	0	1	256	1–3
A5constrad	5	1.5	384	3–5
A0constrad	0	1.5	384	3–5

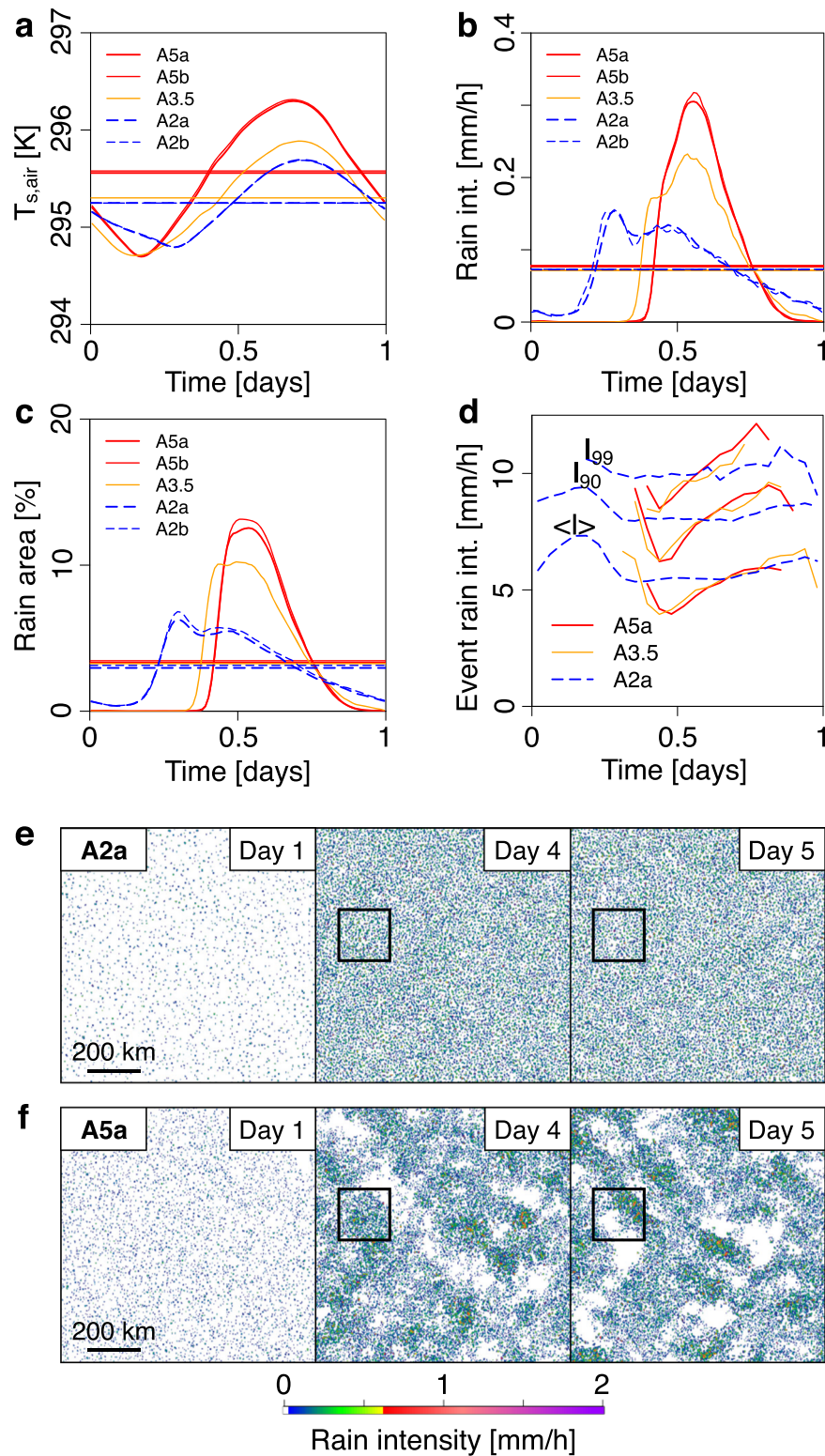
The main four experiments are listed in the first four rows (A5a, A2a, A5b and A2b). The experiments labelled by a star (\*) are equivalent to A2b but constitute an additional, longer-duration run (A2long). All the above experiments were carried out at  $T_s = 298$  K, except A5p2K, which was carried out at  $T_s = 300$  K. A5sea and A2sea were carried out, by assuming a water surface at the lower boundary. In all other experiments, surface evaporation was set to 70% of the value computed for a water surface. A5vent05 and A5vent001 denote experiments, where the ventilation coefficients<sup>52</sup> were reduced to 0.5 and 0.01 of their default values. A5ph3.6h and A5ph6h constitute experiments, where the diurnal cycle of surface temperature was phase-shifted by 3.6 and 6 h. A5constrad and A0constrad are simulations, where the incoming solar radiation was set to a constant equalling its average value.

relatively homogeneous convective activity throughout the domain. During the subsequent model days, convection intensifies for both simulations because near-surface temperatures gradually increase. In A2, the spatial pattern of events remains rather homogeneous (Fig. 1e). In contrast, for A5, an inhomogeneous pattern self-organises, with several subregions receiving pronounced average rainfall, whereas others are all but dry (Fig. 1f, days 4 and 5). Furthermore, for A5, temporal alternations in surface rainfall rate are apparent when comparing one day to the next (compare: Fig. 1f, day 4 vs 5): a cluster on 1 day leaves an almost rain-free area the next day.

To quantify such spatiotemporal inhomogeneities, we determine all surface rain event tracks and compute their centre-of-mass positions. Tracks are thereby defined as spatially and temporally contiguous rainy grid boxes (see “Methods”). We then break the horizontal domain area down into square boxes of side length  $l$ , yielding  $n(l) \equiv (L/l)^2$  such boxes, and determine the number of tracks located in each of the boxes. The probability  $p_l$  of a track occurring within one of the boxes at random would be  $p_l = n(l)^{-1}$  and the binomial

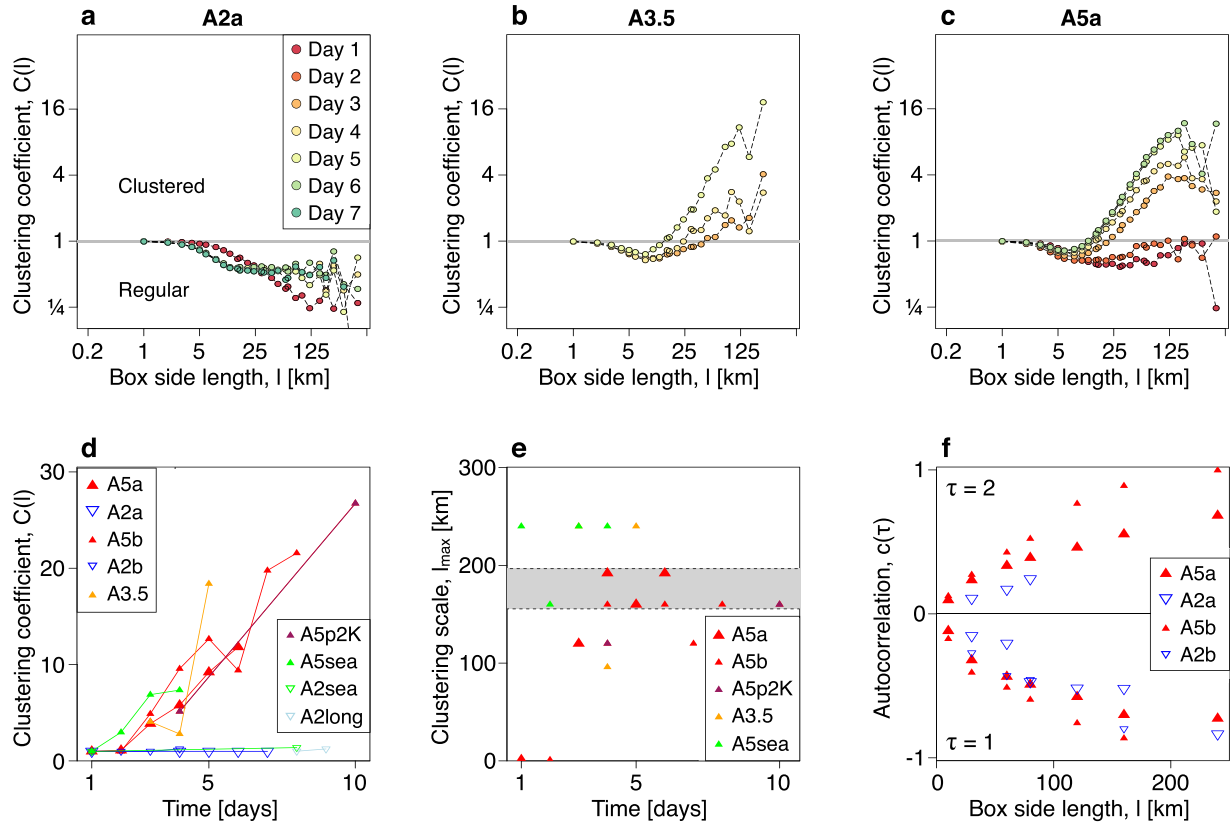
$$P_l(m) \equiv \frac{N!}{(N-m)! m!} p_l^m (1-p_l)^{N-m} \quad (2)$$

hence describes the probability of  $m$  of  $N$  randomly distributed tracks lying in one of these boxes during the model day. The variance of



**Fig. 1 Transition to a clustered rainfall state. a–d** Diurnal cycles of domain-averaged quantities. Each quantity was horizontally averaged for the lowest model level ( $z = 50$  m). Time series represent a compound diurnal cycle, where equal times of day were averaged over all available model days. **a** Temperature for simulations with different imposed diurnal surface temperature amplitudes  $T_{a,i}$  as labelled in the legend. Horizontal lines of corresponding colours represent the time average of each simulation. **b** Analogous to **a**, but for rain intensity. **c** Analogous to **a**, but for rain area fraction. **d** Mean, 90th and 99th percentiles of event rain intensity (conditional on  $I > I_0 = 0.5 \text{ mm h}^{-1}$ ) for A5a and A2a, as labelled (see Table 2 for experiment label details and “Methods” for the definition of rain events). **e** Surface rainfall average during day 1 (spin-up), day 4 and day 5 for A2a. **f** Similar to **e**, but for A5a. Boxes of side length 200 km and scale bars of length 200 km highlight the spatial and temporal variation (see Supplementary Fig. 3).





**Fig. 2 Quantifying the transition to clustering.** **a** Clustering coefficient  $C(l)$  at different box sizes  $l$  for A2a. Curves range from day one (red) to seven (green). The pattern of rain events is regular, as  $C(l) < 1$  throughout. Note the double-logarithmic axis scaling. **b** Analogous to **a**, but for A3.5. **c** Analogous to **a**, but for A5a. In **b** and **c**, at small scales ( $l \sim 10$  km) or early times ( $t < 2$  d), rain events are regularly distributed ( $C(l) < 1$ ), whereas at larger scales ( $l \sim 180$  km) and later times ( $t \geq 3$  d) events are clustered ( $C(l) > 1$ ). **d** Maximum of  $C(l)$  versus time for different simulations (see: legend and Table 1). Note the general increase for large  $T_a$  (A3.5, A5: upward-pointing triangles) but flat behaviour for small  $T_a$  (A2: downward-pointing triangles). **e** Scales of clustering, i.e., the position  $l$  at which the maxima in  $C(l)$  occur in A3.5 and A5. The grey shaded area marks the standard error of  $l_{max}$ , averaged over all times  $t \geq 3$  d. **f** Autocorrelation  $c(\tau)$  for  $\tau = 1$  and  $\tau = 2$  shows increasing  $c(\tau)$  with scale  $l$  for both A2 and A5. Several points for A2 are not shown due to lack of statistical significance at the 1% confidence level. Each data point represents an average over all possible correlations for the experiment at hand: that is, for  $c(2)$ , the pairs (1, 3), (2, 4), etc. were used.

counts  $m$ , given a side length  $l$ , denoted as  $\text{Var}_{\text{ran}}(l; N)$ , is<sup>48</sup>

$$\text{Var}_{\text{ran}}(l; N) = N p_l(1 - p_l), \quad (3)$$

which we compare with the variance of the empirical data

$$\text{Var}_{\text{emp}}(l; N) = \sum_{i=1}^{n(l)} (m_i - \langle m \rangle)^2, \quad (4)$$

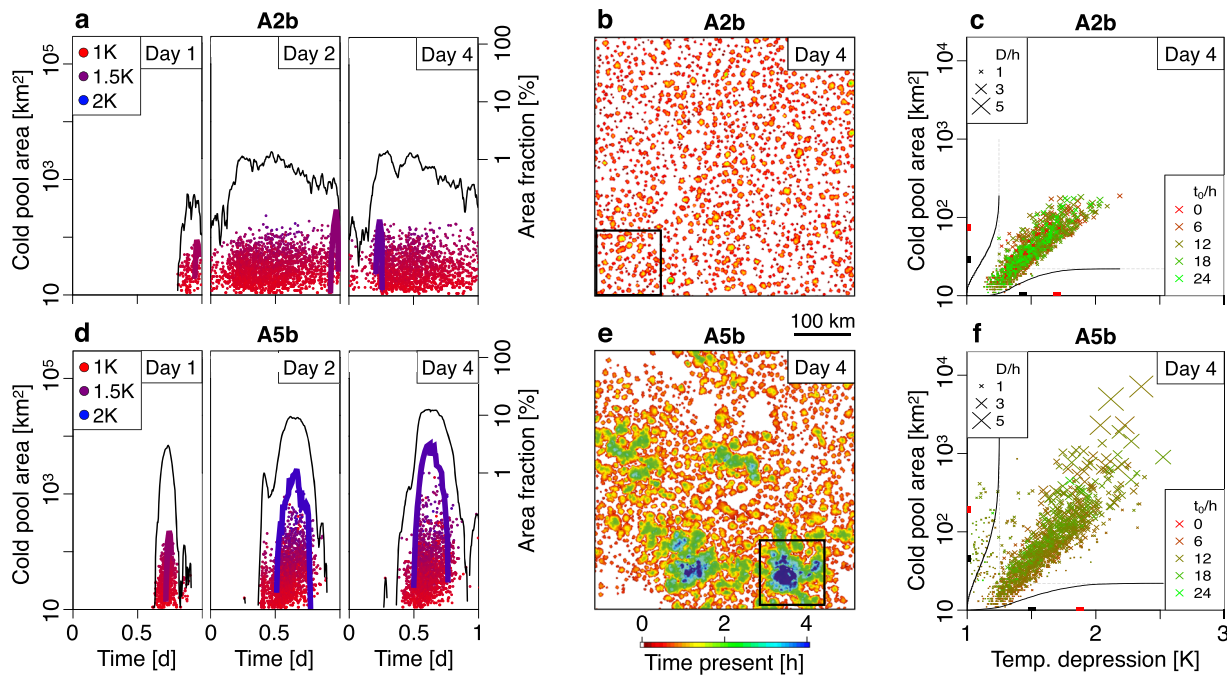
where  $\langle m \rangle \equiv N/n(l)$  is the average number of tracks per box, and the sum runs over all boxes  $i$ . The comparison of the two quantities in Eqs. (3) and (4) hence measures how much variation is obtained empirically when contrasted against completely uncorrelated data. This comparison can be expressed in a clustering coefficient, which we define as  $C(l) \equiv \text{Var}_{\text{emp}}(l; N)/\text{Var}_{\text{ran}}(l; N)$ .  $C(l)$  is below unity when tracks are regularly spaced and above unity when tracks are clustered. The results show that, for small diurnal surface temperature amplitudes (experiment A2a, Fig. 2a)  $C(l) < 1$  for all box sizes  $l$ ; hence, the spacing of tracks is generally more regular than expected at random. For larger diurnal surface temperature amplitudes (A3.5 and A5, Fig. 2b, c) regular spacing with  $C(l) < 1$  is found only for relatively small box sizes of  $l \approx 20$  km, whereas spacing at larger box sizes is strongly clustered, that is,  $C(l) \gg 1$ . Moreover, this clustering increases from day to day (Fig. 2d).

We also define a box size  $l_{\text{max}}$  at which  $C(l)$  is maximal. Despite some variation,  $l_{\text{max}} \approx 180$  km can be identified from A5 (Fig. 2c, e). In addition, we measure the autocorrelation  $c(\tau)$  between day  $d$  and day  $d + \tau$  by computing the Pearson correlation coefficient of

daily mean precipitation rates from all grid boxes (Fig. 2f). We find rainfall to be anticorrelated from one day to the next for all box sizes ( $c(1) < 0$ ), suggesting a local inhibitory effect of rain and positively correlated two days into the future ( $c(2) > 0$ ). The magnitudes of  $c(1)$  and  $c(2)$  both increase with box size  $l$ , but appear to level off near  $l = 200$  km (Fig. 2f).

#### Clustering as a result of cell-density difference

We now explore the mechanistic origin of the clustering. Differences in the areal number density, that is, the count of rain cells per unit area, between A2 and A5, are evident from the rainfall diurnal cycle (Fig. 1b, c). The peak in rain intensity and fractional area covered by rain cells is nearly twice as high for A5 compared with A2. Given that the area, intensity and duration of each individual rain cell are all similar for A2 and A5 (Supplementary Table 1), roughly twice the number of rain events will coincide during the peak of rainfall for A5 compared with the peak of A2. As mentioned, CPs are known to mediate interactions between convective rain cells<sup>32,49</sup>. We expect interactions to become more relevant at increased rain cell density because CPs will more often collide as they spread along the surface. To quantify density effects on the clustering dynamics in A2 versus A5, we first perform a simple CP tracking. The method identifies spatially and temporally contiguous patches of low buoyancy, measured by a 1 K temperature depression (for detail, see “Methods”). In A2, CP areas typically do not exceed 500 km<sup>2</sup>, have modest temperature



**Fig. 3 Cold pool merging and deepening.** **a** CP occurrence time, maximum area and average temperature depression (colours from red to blue, see legend) on days 1, 2 and 4 of the experiment A2b. The black curve indicates the total CP area at each time, whereas the thick coloured curve (compare legend for colours) highlights the time series of the largest CP during the particular day. **b** Areas covered by CPs during day 4 of A2b, the colours (see colour bar) indicate the duration during which CPs were present (compare: Supplementary Fig. 5). The scale bar is 100 km long. **c** CP area versus the corresponding maximum of areal mean temperature depression (day 4). Symbol sizes indicate CP lifetime and colours indicate occurrence time within the model day (see legends). Black and red rectangles along the axes indicate the median and 90th percentile of the corresponding quantity. Note the logarithmic vertical axis scale in **a** and **c**. **d–f** Analogous to **a–c**, but for A5b.

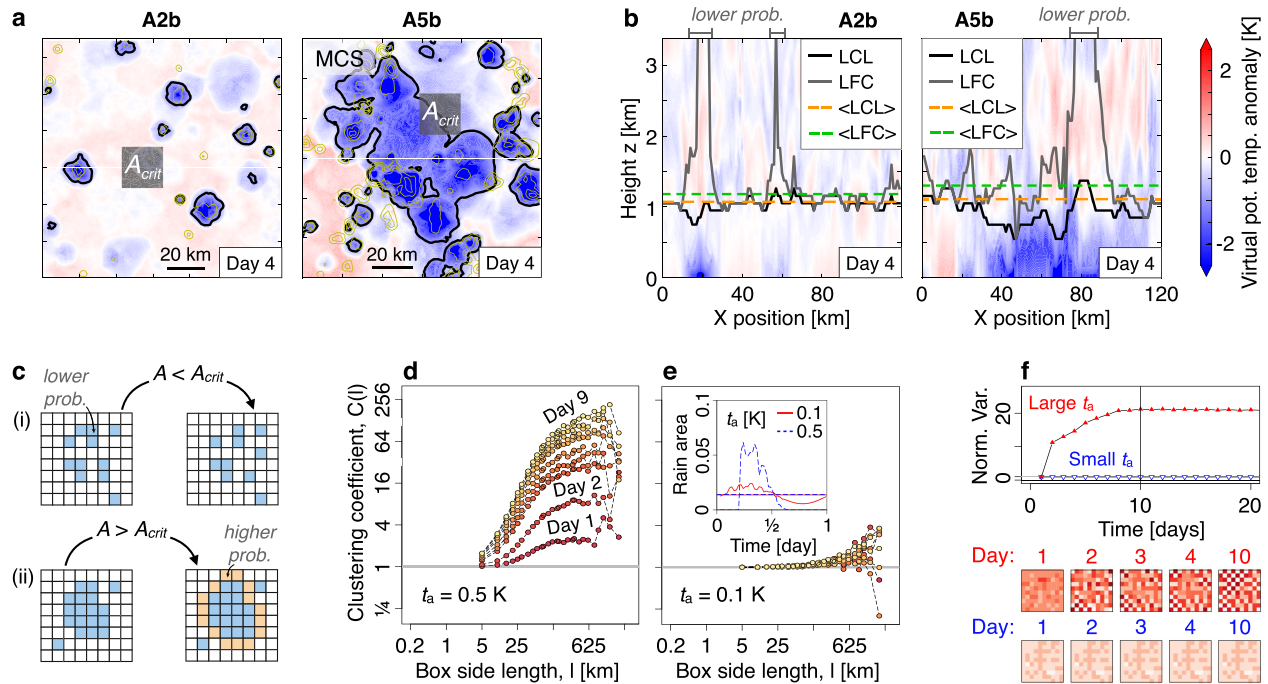
depressions and lifetimes of generally less than 2 h (Fig. 3a–c). In A5, where CPs only occur during parts of the day, CP areas often exceed  $10^3 \text{ km}^2$ , and such large CPs have much stronger temperature depressions and substantially longer lifetimes (Fig. 3d–f). The formation of very large and strongly negatively buoyant CPs suggests that CPs from distinct rain cells often bunch together before each CP has fully expanded. We employ the CP tracking to detect merging events, that is, instances where two previously separate CP areas combine. For A5b, CP merging is indeed ubiquitous, whereas it is all but absent in A2b (Supplementary Fig. 4c).

In addition, we compute the CP height, measured by detecting temperature anomalies in the vertical dimension. In A5b, when CPs first appear, they show heights comparable to those of A2b (Supplementary Figs. 4a, b and 5). Subsequently, a double peak forms, where groups of merged CPs obtain much larger heights, reaching close to the level of free convection ( $\approx 1100 \text{ m}$ ). Larger CP heights  $h$  and deeper temperature depressions  $T'$  are consistent with larger mean surface wind speed  $v_{cp} \sim (hT')^{-1/2}$  under CPs<sup>50</sup> (compare: panels in Supplementary Fig. 5) and surface fluxes (Supplementary Fig. 1e–h), resonating with markedly enhanced near-surface moisture and convective available potential energy (CAPE) along CP boundaries in A5 (compare: Supplementary Fig. 6e, j).

We aim to formulate a simplified model that incorporates the observed number-density differences. To this end, consider first a detailed comparison of CPs formed in A2 and A5: in A2, CPs are spatially isolated from one another, and the area covered by each CP remains small (Fig. 4a). In A5, many CPs occur so close to each other, that their temperature anomalies inevitably merge (Fig. 4a; Supplementary Fig. 4c), forming a large patch of dense air. This combined CP, which we associate with the emergent MCS, reaches a greater height and develops a relatively cold and dry

region in its interior where convection is suppressed. This inhibitory effect in the MCS interior can be ascribed to the divergence of the level of free convection (LFC) (Fig. 4b) and a depletion of CAPE (Supplementary Fig. 6e, j).

In contrast, as the combined CP spreads outward from the dense region of rain cells, new cells are often triggered at its front—further feeding the MCS. At this combined CP's gust front, hence at the periphery of the emergent MCS, the greater CP height allows boundary-layer environmental air to be forced higher up and set off new rain cells. In addition, the CP's surroundings benefit from the moisture transported by the combined CP's front and the additional latent heat flux and increased CAPE along the combined CP's edges (compare Supplementary Fig. 6e, j). Indeed, many subsequent rain cells do form near the perimeter of the combined CP, whereas this is rare for A2 (thin black contours in Fig. 4a). To quantify this moisture redistribution from within the MCS to its surroundings, we contrast domain subregions of A5, which receive intense versus weak precipitation during a given model day (Supplementary Fig. 6f–h). Regions of intense rainfall are characterised by enhanced moisture near the cloud base ( $z \sim 1 \text{ km}$ ) before precipitation onset, but marked depletion after rain has occurred. Conversely, areas of weak rainfall show nearly a “mirror image”, with depressed moisture before but enhanced values after precipitation. The bi-diurnal dynamics for A5 can hence be characterised as an alternation of cloud-base moisture, driven by the lateral expansion of MCSs, which leaves behind an inhibitory dry patch in their interior that is marked by low CAPE (Supplementary Fig. 6i, j). On the subsequent day, this drying in locations of previous MCSs leads to relatively enhanced near-surface and cloud-base moisture in those regions not affected by these previous MCSs—in turn providing favourable conditions for convective activity there (for detail, see Supplementary Information Text and Supplementary



**Fig. 4 A simplified model for convective clustering.** **a** Examples of typical CPs on day 4 of A2b (15:35) and A5b (14:00) showing virtual potential temperature anomaly at  $z = 50$  m (compare: boxes in Fig. 3b, e for context). Thin yellow lines show accumulated surface rainfall (1-, 5- and 10-mm contours) within the subsequent hour. A scale bar and an area  $A_{crit}$  are marked (compare: panel c). CP areas (bold black contours of  $-1$  K anomalies) exceeding  $A_{crit}$  are rare in A2b, but frequent in A5b. The combined CPs in A5b can excite subsequent rainfall and feed the emergent MCS. **b** Respective  $x$ - $z$  cross-sections along the white horizontal lines in **a** including the lifting condensation level (LCL) and level of free convection (LFC), and corresponding domain means (dashed). Marked regions correspond to lowered triggering probability (compare: panel c). **c** Schematic for the simplified model dynamics. (i) Low-density subdomain with vacant (white) and active sites (blue). (ii) Similar to (i) but for high density, showing also boundary sites (orange) with higher probability. **d**  $C(l)$  for  $t_a = 0.5$  K, analogous to Fig. 2a–c but now for the simplified model (days as marked). Note the double-logarithmic axis scaling. **e**, Analogous to **d**, but for smaller  $t_a$ . Inset: diurnal cycle for rain area for the simplified model for two values of  $t_a$  (compare: Fig. 1b–d). **f** Simple checkerboard model, explaining the increase in variance over time for large  $t_a$  (red symbols) and small  $t_a$  (blue symbols). The grids below show the patterns at days 1–4, 10 for large and small  $t_a$ , respectively. Box colours white  $\rightarrow$  faint red  $\rightarrow$  red  $\rightarrow$  dark red indicate increasing density.

Fig. 6). We further find a positive correlation between cloud-base moisture and precipitation, which could help promote the transition from shallow to deep convection<sup>29,51</sup> and potentially reinforce the updraughts. In A2, day-to-day moisture alternations are lacking (see Supplementary Fig. 6g, h for A5b to Supplementary Fig. 6b, c for A2b), a finding that falls in line with the absence of organised convection. As with rainfall, the spatial pattern of cloud-base moisture again reflects a coarsening of scales for A5 relative to A2 (Supplementary Fig. 6d, i).

In total, the analysis suggests that an MCS emerges, when several CPs combine, become deeper and excite new rain cells along their common gust front. The emergence of an MCS on one day causes a suppressed region the subsequent day. As a further check, we weaken CPs by decreasing the ventilation coefficients ( $a_v$  and  $b_v$  in Eq. (24) of ref.<sup>52</sup>) to a fraction of their default values. These coefficients control rain evaporation, hence temperature depression  $T'$  and CP propagation ( $\sim T'^{1/2}$ ) (for details, see “Methods”). Indeed, decreasing these coefficients systematically reduces the spatial extent of the rain-free regions (Supplementary Fig. 3).

#### Clustering from a two-level atmosphere model

A key characteristic of A5 is that parts of the day see no rainfall at all. In contrast, immediately after the onset of rain (near mid-day, Fig. 1b), the area covered by precipitation is relatively large—corresponding to a high number density of rain events and CPs. Our simplified model describes the system domain as a square lattice of sites, where each site can reside in two different states. For simplicity, we consider each pair of a rain cell and its CP as one

entity. We let this entity occupy one lattice site, termed an “active site”, and say that it fills the area of a rain cell  $a_0 \approx 25$  km<sup>2</sup> (Supplementary Table 1). Lattice sites not active are considered “vacant”. In practice, our model mimics the qualitative clustering dynamics by allowing MCSs to emerge, whenever a sufficient number of active sites occur close to each other.

Assume that, at a given time, the fraction of active sites is  $p_0$ , and sites are independently populated. That is, each site of a square lattice contains a rain event at probability  $p_0$ . To exemplify, in Fig. 4c, i,  $p_0$  is relatively small, and few sites are active. In Fig. 4c, ii,  $p_0$  is larger, more sites are therefore active and many of them are now neighbours. To incorporate the effect of an MCS, we simply let the remaining vacant sites be more likely to become active when they are immediate neighbours to a sufficiently large area  $A > A_{crit}$  of spatially contiguous active sites (compare: Fig. 4a). This is accomplished by assigning increased probabilities to the neighbourhood sites (Fig. 4c, ii, orange sites). When  $p_0$  is small ( $p_0 \ll 1$ ), the system will, however, be unlikely to contain contiguous rain areas exceeding  $A_{crit}$  (compare: shaded box in Fig. 4a). To be explicit: the probability of finding two active sites on two neighbouring sites is proportional to  $p_0^2$ , and this probability will decay exponentially with the number of sites contained in the contiguous area<sup>53</sup>.

But how does  $p_0$  emerge from the diurnal cycle dynamics, and how can bi-diurnal temporal correlations be captured (Fig. 2)? To self-consistently incorporate these features, we describe the population of rain cells within a two-layer atmosphere model consisting of a prescribed boundary-layer temperature  $T_{bl}$ , varying sinusoidally with a small diurnal surface temperature amplitude



$t_a$  (see Fig. 1a), and an interactive free-tropospheric temperature  $T_{ft}$ . We compute the probability for a rain cell to become active by coupling it to the temperature difference between the boundary layer and the free troposphere  $\Delta T = T_{bl} - T_{ft}$ , as an approximation for atmospheric stability. When  $T_{bl}$  increases during the model day, rain events will eventually be set off. Two key processes impact on the free-tropospheric temperature: thermal radiation to space reduces  $T_{ft}$ , whereas latent heat transfer by rain events increases it. The boundary-layer temperature ( $T_{bl}$ ) is not directly affected by rain cells. However, buoyancy depressions, arising mainly from sustained drying after rain events, are implemented by an inhibitory potential. In practice, once an active site transitions back to vacant, it needs to “wait” until it can become active again (for details, see “Methods”).

We implement reasonable coefficients for these processes and find the simulations to reach a repetitive diurnal cycle (Fig. 4e, inset). Indeed, for small  $t_a$ , modest rainfall is present during the entire day, whereas for larger  $t_a$ , rainfall is either intense or absent. Time-averaged rain areas for large and small  $t_a$  match (compare: Fig. 1b), a result of the radiative constraint and in agreement with the numerical experiments. The earlier phase in Fig. 4e, inset versus Fig. 1b, can be attributed to the fact that we directly prescribed atmospheric boundary-layer temperature  $T_{bl}$ , rather than surface temperature itself, causing an immediate precipitation response.

Considering the variance of the spatial pattern for large  $t_a$ , the simplified model indeed produces increased clustering over time and larger spatial scales (Fig. 4d, colours from red to green, and Supplementary Fig. 7). Conversely, clustering is absent for small  $t_a$  (Fig. 4e). This can be explained intuitively: when the thermal forcing caused by the boundary-layer temperature  $T_{bl}$  increases rapidly during the day, many rain events will be set off during a short time—leading to large  $p_0$  during those times. The negative feedback on the free-tropospheric temperature  $T_{ft}$  will then rapidly cause the “budget” of rainfall to be used up. MCS will form as long as the increased probability at the edges of the contiguous CP patches counteracts the ongoing increase of  $T_{ft}$ . Hence, MCSs will be able to spread, as long as this is the case, thus setting the time ( $\approx 6$  h) and space ( $\approx 100$  km) scale for MCS, which is significantly larger than the scale of a single rain event ( $\approx 1$  h and  $\approx 5$  km).

### Conceptualising further

The two scales of individual rain cells ( $\sim 5$  km) and MCS ( $\sim 100$  km) allow for a further simplified conceptual view, which builds on domain total moisture being conserved from day-to-day. In this view, the occurrence of rain cells is directly proportional to the available moisture and rain cells act to redistribute moisture, once they form an MCS. Take the model domain to be broken down into a square lattice consisting of blocks. Each “block” consists of  $20 \times 20$  sites for single rain cells of area  $a_0$ , together yielding a block size of  $100 \text{ km} \times 100 \text{ km}$ , which provides space for an entire MCS. In each block, an MCS is only set off if a sufficient number of rain cells are present. We find that a simple set of rules can model the MCS dynamics: (1) first, initialise each block in this square lattice, by assigning an integer, encoding the number of rain cells, drawn from a binomial distribution; (2) update the system once per day, by letting all blocks above a particular threshold (an MCS forms) hand over their content (the moisture transported by the MCS) at equal parts to their four neighbouring blocks; after this distribution step, all sites collect the spilt content, hence conserving total moisture.

When the mean of the binomial distribution is low compared with the threshold, which is the case for small  $t_a$ , no MCS will form and no redistribution will take place (Fig. 4f, blue points, and lower row of squares). In contrast, when  $t_a$  is sufficiently large, a long sequence of reallocations will occur, leading to a checkerboard-

like clustering, which strengthens in time (red curve and upper row of squares). This example is sufficient to capture the increase of normalised variance for A5 and the lack of it for A2 (Fig. 4f).

### DISCUSSION

Our study analyses the spontaneous emergence of convective clustering, when the amplitude of the diurnal cycle surface temperature is varied. This “diurnal self-aggregation” occurs without imposing spatial scales by external forcing, as with radiative–convective equilibrium (RCE) self-aggregation<sup>8–10</sup>. Clearly, the two cases differ regarding the surface boundary conditions. However, also the spatial and temporal scales of the clustering, and in particular the physical processes causing it, differ. In RCE self-aggregation, clusters are temporally coherent over multiple days, whereas in diurnal self-aggregation, the spatial patterns of convection are almost opposite from one day to the next. The findings here suggest that, when the diurnal surface temperature variation is sufficiently large, the clustering, and thereby the potential for flash floods, increases from day-to-day. Increasing model horizontal resolution to 0.5 or 0.2 km (A5c and A5d in Supplementary Fig. 3) leads to even stronger clustering effects, in agreement with recent findings for continental surface conditions<sup>42</sup>. It should be tested if higher resolution leads to stronger cold-pool interaction, hence more activation in areas with dense rain cell occurrence. The current findings suggest that observational studies on extremes<sup>54–56</sup> should additionally focus on diurnal surface temperature amplitude, not just the mean.

In RCE self-aggregation, the standard explanation for clustering invokes arguments based on differential cooling and circulation changes induced by this cooling. Remarkably, the memory enabling clustering in diurnal self-aggregation is purely thermodynamically driven: we do not detect any sustained changes of circulation that harbour a memory from one day to the next (Supplementary Fig. 1i, j). The findings hence suggest two possible “shortcuts” to observed mesoscale clustering over the ocean: (i) at times, sea-surface temperature oscillations may be sufficient to kick-start clustering; (ii) clustering may emerge over land surfaces, where a strong diurnal cycle prevails, and may then be advected over the sea, where it could grow further under more RCE-like conditions. Practically, it could be checked through simulations, whether a state of persistent, spatially and temporally coherent self-aggregation is reached more quickly when sea-surface temperatures transiently oscillate—thereby exciting initial clustering on the scale of  $l_{\text{max}}$ . Both “shortcuts” require that clustering can persist or grow once formed—a requirement consistent with hysteresis effects<sup>57</sup>. In our conceptual model, hysteresis is, in fact, easy to achieve. Once an imbalance between neighbouring grid cells is established, sustained re-distributions will be possible, even when the overall event number density is then lowered. It should further be explored, if the day-to-day variation we find can be amplified by an interactive surface, such as by “cloud shading”<sup>8,13</sup>.

The specific feature of the MCS, formed by higher cell and thus cold-pool density at larger diurnal surface temperature amplitude, is to trigger new convective cells at its periphery. These new cells form cold pools, feeding the emergent MCS and further forcing updraughts near its boundary. Notably, new convective cells are also formed within the interior of the MCS (Fig. 4a), a finding in line with collision effects of multiple cold-pool gust fronts<sup>17,32</sup>. These interior cells could further act to deepen and cool the combined cold pool driving the MCS expansion. Together, MCSs hence act to excite new convection both within and around the combined cold-pool area. In conventional self-aggregation studies, strong cold pools were found to inhibit clustering<sup>16</sup>. However, under the diurnal cycle conditions studied here, they combine to promote the formation of MCSs due to the high numbers of cold pools occurring in spatially confined subregions within a relatively short time.



Altogether, the present work interpolates between the established RCE set-up for oceanic convection and typical boundary conditions for continental convection. Over the ocean, the origin of initial clustering could be found in modest sea-surface temperature fluctuations—in line with those observed<sup>18–20</sup>. Over continents, fluctuating surface temperatures are well-known to imprint pronounced diurnal rainfall variations<sup>31,34–38</sup>, and observed MCS rainfall was reported to increase<sup>2,58</sup>, elevating the risk of flash floods. In summary, the present results suggest that explanations for tropical and summertime mid-latitude flash floods might also be found in temperature variations.

## METHODS

### Large-eddy model, boundary and initial conditions

We simulate the convective atmosphere using the University of California, Los Angeles (UCLA) Large Eddy Simulator (LES) with sub-grid-scale turbulence parameterised after Smagorinsky<sup>59</sup>. This is combined with a delta four-stream radiation scheme<sup>60</sup> and a two-moment cloud micro-physics scheme<sup>61</sup>. Radiation interacts with the atmosphere including clouds, but it does not impact the surface temperature and hence surface fluxes. The diurnal insolation cycle is chosen typical for the equator (Supplementary Fig. 2d). Rain evaporation is implemented after Seifert and Beheng<sup>52</sup>. Diurnally oscillating, spatially homogeneous, surface temperature,  $T_s(t)$ , is prescribed, as

$$T_s(t) = \bar{T}_s - T_a \cos(2\pi t/t_0), \quad (5)$$

where  $\bar{T}_s = 298$  K,  $t_0 = 24$  h is the duration of the simulated model day,  $\bar{T}_s$  represents the temporal average and  $T_a$  the amplitude of  $T_s(t)$ . Surface heat fluxes are computed interactively and depend on the vertical temperature and humidity gradients as well as horizontal wind speed. Horizontal wind speed is approximated using the Monin–Obukhov similarity theory, which applies scaling arguments to assume logarithmically increasing horizontal wind speed with height<sup>62</sup>. Temperature and humidity were initialised using observed profiles that potentially represent convective conditions<sup>37</sup>. However, due to the repeated diurnal cycle forcing, the system eventually establishes a self-consistent vertical temperature and moisture profile.

### Model grid, dynamics and output

The model integrates the anelastic equations of motion on a regular horizontal domain with varying horizontal grid spacing  $dx$  and periodic boundary conditions (Table 1). The vertical model resolution is 100 m below 1 km, stretches to 200 m near 6 km and reaches 400 m in the upper layers, with the model top located at 16.5 km. A sponge layer is employed between 12.3 km and the model top. Horizontal resolution  $dx$  and domain size vary (Table 1). The Coriolis force and the mean wind were set to zero with weak random initial perturbations added as noise to break complete spatial symmetry. For all two- and three-dimensional model variables, the output timestep varies between experiments between  $\Delta t_{\text{out}} = 5$  min and 15 min. At each output timestep, instantaneous surface precipitation intensity, as well as the three-dimensional moisture and velocity fields, is recorded for the entire model domain. In addition, at 30-s and 5-min intervals, respectively, spatially as well as horizontally averaged time series were extracted from the numerical experiments.

### Sensitivity experiments

The principal focus of this study is the response to different values of diurnal surface temperature amplitude,  $T_a$ . The main experiments (A5a, A2a, A5b and A2b) were carried out at  $\bar{T}_s = 298$  K, contrasted  $T_a = 2$  K and  $T_a = 5$  K (Table 1) and used  $dx = 1$  km horizontal model resolution. One intermediate value of  $T_a = 3.5$  K was tested to constrain the transition to clustering further. The control case of  $T_a = 0$  K was also tested as a benchmark. These simulations assumed that surface latent heat fluxes occurred at 70% of their potential value. This mimics a land surface, which is reasonable given the mean surface latent and sensible heat fluxes of  $LHF \approx 57$  W/m<sup>2</sup> and  $SHF \approx 18$  W/m<sup>2</sup>, respectively, yielding a Bowen ratio of  $B \approx 0.30$ , realistic for forested land. These main experiments already explore domain-size effects, as the clustering observed could be influenced by the finite system size. We find that the results vary little between these domain sizes. We conducted additional experiments to test the sensitivity to a range of modifications: horizontal model resolution (A5c, A5d, Supplementary Fig. 3), where  $dx = 0.5$  km and  $dx = 0.2$  km were used; cold-pool

strength (A5vent05, A5vent001, Supplementary Fig. 3), where the ventilation coefficients  $a_v$  and  $b_v$  in Eq. (24) of Seifert and Beheng<sup>52</sup> were respectively reduced to 0.5 and 0.01 of their default values; surface conditions, where the surface evaporation was that of a sea surface (Supplementary Fig. 2a–c, as well as A5sea and A2sea in Supplementary Fig. 3, here  $B \approx 0.15$ ); the phase of the surface temperature diurnal cycle, described by a term  $\tau$ , when generalising Eq. (5) to

$$T_s(t) = \bar{T}_s - T_a \cos(2\pi(t - \tau)/t_0) \quad (6)$$

relative to that of the diurnal insolation cycle.  $\tau = 3.6$  h and  $\tau = 6$  h (A5ph3.6h and A5ph6h, respectively) were applied, accounting for a possible lag of the surface temperature peak towards the mid or late afternoon, as well as a constant solar radiative forcing (A0constrad and A5constrad). The dominant effect of nonzero  $\tau$  is to shift the diurnal precipitation cycle accordingly (Supplementary Fig. 2e), whereas clustering still occurs (Supplementary Fig. 3). In contrast, setting insolation to the temporal average (dashed line in Supplementary Fig. 2d) has the effect of shifting the diurnal cycle of precipitation to earlier times (Supplementary Fig. 2e, dotted vs solid red line). However, also for constant insolation, clustering is still observed for A5 (Supplementary Fig. 3) (for further details, see Table 1 and Supplementary Table 1).

### Rain cell and cold-pool tracking

In all experiments, we track rain cells using the Iterative Rain Cell Tracking Method (IRT<sup>63</sup>). In the two-dimensional surface precipitation field corresponding to any output timestep, the IRT first detects all spatially contiguous patches (using connectivity 4) of rain intensities exceeding a threshold  $l_0 = 0.5$  mm h<sup>−1</sup>, termed *rain objects*. Tracks are then identified, by determining any rain objects that overlap from one output timestep to the next. In the case of merging of tracks, the track of a larger area is continued under the same track ID (the threshold value  $\theta$  was set to unity). In our variance analysis, we use the set of coordinates formed by the initial positions of each track, which is defined here as the precipitation-weighted centre of mass of the first rain object belonging to each track. The locations of all subsequent rain objects of the same track are discarded in our variance analysis, to avoid artefacts from double counting of the positions of the nearly collocated objects.

To account for possible changes in the CP extent between the different simulations, CPs were tracked by a simple temperature-depression method. In this method, the IRT was modified, such that at any timestep, any grid box within the lowest model level with a temperature depression, measured relative to the spatial mean at the same time, exceeding a threshold of one Kelvin, was recorded. As for rainfall, spatially contiguous patches of temperature depression were collected into objects and assigned a unique track identifier. To increase the signal, only objects exceeding an area threshold of 10 km<sup>2</sup> were considered. Going forward in time, overlapping objects were identified and considered to be the same CP track. Again, the track of the larger CP area is continued under the same track identifier<sup>63</sup>.

For CP height, a criterion based on a fixed temperature deviation was found inappropriate, as temperature differences are smaller aloft. Therefore, we measure CP height by evaluating the average temperature of the domain and the standard deviation at each output timestep and vertical model level  $k$  and comparing the local temperature  $T_0$  at each grid cell  $(i, j, k)$  with this value. If  $T_0$  is two standard deviations below the domain mean at the same height level  $k$ , this grid box is considered part of a CP. Within each column  $(i, j)$ , the CP height is determined as the highest level that fulfils this criterion.

### A simplified mesoscale convective system (MCS) model

A square lattice of  $s \times s$  sites is initialised, where the area of each site is taken as the average area  $a_0$  occupied by a single convective rain cell,  $a_0 \approx 5$  km  $\times$  5 km (compare: Supplementary Table 1). The total domain area  $A$  hence is  $A \equiv a_0 s^2$ . Each site can be in one of two convective states, namely active or inactive. Our model assumes boundary-layer processes to be local. In contrast, the free troposphere acts as a reservoir of large heat capacity, where heat is quickly redistributed in the entire free troposphere by gravity waves<sup>64</sup>; thus, spatial homogeneity is assumed there. The model incorporates three fundamental processes affecting event formation: (1) spontaneous activation due to the moderate drive of the diurnal cycle temperature forcing, (2) refractory dynamics due to CPs and (3) strong activation due to combined CPs ahead of MCS fronts. Below, we discuss

**Table 2.** Parameters in the simplified MCS model.

Model parameter	Numerical value	Description
$a_0$	25 km <sup>2</sup>	Area of a single site
$A_{\text{crit}}$	500 km <sup>2</sup>	Critical area for MCS formation
$p_{\text{act}}$	0.8	Activation potential along MCS fronts
$\tau_{\text{cp}}$	3 h	Duration when CPs can activate new rain cells
$p_{\text{inh}}$	−0.8	Refractory potential underneath cold pools
$\tau_{\text{inh}}$	24 h	Humidity inhibitory time
$t_a$	0.1 K, 0.5 K	Diurnal near-surface temperature amplitudes

the estimation of model parameters for these processes (summarised in Table 2).

**Spontaneous activation.** To define a vertical temperature gradient, we consider quantities  $T_{bl}$  and  $T_{ft}$ , which represent the deviation of the boundary layer and free-tropospheric temperature from their assumed steady-state values.  $T_{bl}(t)$  is prescribed and oscillates harmonically as

$$T_{bl}(t) = -t_a \cos(2\pi t/t_0), \quad (7)$$

where  $t_0 = 24$  h. The free-troposphere temperature anomaly  $T_{ft}$  is initialised to zero. We also define the temperature difference  $\Delta T \equiv T_{bl} - T_{ft}$ , and its normalised version  $\Delta T' \equiv \Delta T/\Delta T_0$ , where the reference scale  $\Delta T_0 \approx 0.3$  K is taken as two standard deviations of LES-simulated near-surface temperature.  $\Delta T'$  serves as a proxy for both convective available potential energy (CAPE) and convective inhibition (CIN). The basic dynamics proceeds in discrete model timesteps of 0.5 h, reflecting the typical timescale of convective cloud formation. At each timestep, the default probability  $p_0$  for a vacant site to become active is taken as

$$p_0 = \begin{cases} 0 & \text{if } \Delta T' < 0 \\ \Delta T' & \text{if } 0 < \Delta T' < 1 \\ 1 & \text{if } \Delta T' > 1. \end{cases} \quad (8)$$

Equation (8) makes the qualitative assumption that there is no activity at all for stratified conditions ( $\Delta T' < 0$ ), and events occur with certainty when the temperature gradient is much larger than typical fluctuations. Both of these limits could be softened, as could the assumption of linearity at intermediate  $\Delta T'$ . One could equally argue for a smooth function of  $\Delta T'$ , such as an error function. Nonetheless, Eq. (8) captures the observed fact that more initial activity occurs when surface temperature changes quickly. Without any further perturbations, the probability  $p_{ij}$  to initiate convection in a site  $(i, j)$  will be equal for all sites ( $p_{ij} = p_0$ ).

**Spatial structure.** Notably, Eq. (8) does not depend on the position within the lattice. Local modifications cause spatial structure (compare: Fig. 4a). Cold pools have two effects: reduction of local temperature and reduction of local humidity. However, the recovery timescale for the former is fast ( $\tau_{\text{cp}} \approx 3$  h), whereas that of the latter can be slower ( $\tau_{\text{inh}} \approx 24$  h, compare: Supplementary Fig. 6)<sup>27</sup>. To consider the former, we take active sites to persist for  $\tau_{\text{cp}}$ , during which no further rain cell initiation is possible at the same site. Simultaneously, the local probability  $p_{ij}$  is reduced as  $p_{ij} = p_0 - p_{\text{inh}} \exp(-\delta t/\tau_{\text{inh}})$ , where  $\delta t$  is the time after the occurrence of the rain event and  $p_{\text{inh}} > 0$ .  $p_{\text{inh}}$  controls the fidelity of the anticorrelation from day-to-day, and the results are qualitatively not affected by it.

An MCS is defined to occur when a 4-connected contiguous patch of active sites exceeds the threshold area  $A_{\text{crit}} \equiv n_0 a_0$ , with  $n_0 = 20$  (Fig. 4). While this is the case, and  $\Delta T' > 0$ , the probability  $p$  at each of the surrounding sites  $(i, j)$  becomes  $p_{ij} \rightarrow p_{ij} + p_{\text{act}}$ , which acts to lower the barrier presented by CIN. The increased CIN above CPs is reflected in the strongly elevated LFC above the CP or MCS (Fig. 4a, b). When  $p_{ij}$  is no longer a surrounding site, the term  $p_{\text{act}}$  is no longer applied for this site. The initiation probability  $p_{ij}$  is then defined analogously to the one in Eq. (8).

**Radiative constraint.** Generally,  $T_{bl}$  rises during the day as the surface is heated by insolation.  $p_{ij}$  will then eventually become positive at some sites, and rain cells can be produced there. When this occurs, latent heat is transferred to the free troposphere, increasing  $T_{ft}$ , thus subsequently lowering  $\Delta T$  and thereby  $p_{ij}$ . Simultaneously, the site  $(i, j)$  is shut down by

the adverse buoyancy effects through CP formation.  $T_{ft}$  will relax by heat loss ( $P_{\text{out}} \approx 200 \text{ W m}^{-2}$ ) through outgoing thermal radiation. According to the Stefan–Boltzmann law,  $P_{\text{out}} = \sigma T_{\text{eff}}^4$ , with  $\sigma \approx 5.7 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ , the Stefan–Boltzmann constant—translating to an effective emission temperature  $T_{\text{eff}} \approx 250$  K. Through  $P_{\text{out}}$ , the free troposphere would hence cool by  $\sim 2 \text{ K d}^{-1}$ . As a crude estimate of the change in  $T_{ft}$  through latent heat transfer, we estimate the free-tropospheric heat capacity  $C_{ft} = M_{ft} c_{pd}$ , assuming dry air.  $c_{pd} \approx 1 \text{ kJ kg}^{-1}$  and the mass of the free troposphere

$$M_{ft} = \int_{z=z_{\text{LCL}}}^{z_{\text{top}}} dz \rho(z) \approx 8 \times 10^3 \text{ kg}, \quad (9)$$

using  $z_{\text{LCL}} \approx 1$  km and  $z_{\text{top}} \approx 16$  km for the lifting condensation level (LCL) and top of the troposphere, respectively. Hence,  $C_{ft} \approx 8 \times 10^6 \text{ J m}^{-2} \text{ K}^{-1}$ . From rain cell tracking of our LES simulations, we obtain the average rain cell lifetime to be  $\approx 1$  h (Supplementary Table 1), the mean cell area  $a_0 \approx 20 \text{ km}^2$  and the average cell rain rate of  $\approx 4 \text{ mm h}^{-1}$  (Supplementary Table 1); hence, each rain cell can be assumed to yield  $M_{\text{event}} = 4 \text{ kg m}^{-2}$  of liquid water. Assuming that the corresponding latent heat was previously deposited in the free troposphere upon ascent, each rain event heats the free troposphere by  $Q_{\text{event}} = L_v M_{\text{event}}$ . We take this heat to increase  $T_{\text{eff}}$  accordingly by  $\delta T_{\text{eff}} = Q_{\text{event}} C_{ft}^{-1} \text{ s}^{-2}$ , where  $\text{s}^{-2}$  results from the ratio  $a_0/A$ , yielding  $\delta T_{\text{eff}} \approx 1 \text{ K}$ : if one rain event occurred instantaneously at each site, the free-tropospheric temperature would rise by 1 K.

## DATA AVAILABILITY

The rainfall simulation data that support the findings of this study are available in the online repository with the identifier 10.5281/zenodo.3839932.

## CODE AVAILABILITY

The computer code to generate the main analysis results is available in the online repository with the identifier 10.5281/zenodo.3841916.

Received: 2 February 2020; Accepted: 7 July 2020;

Published online: 30 July 2020

## REFERENCES

- Moncrieff, M. W., Liu, C. & Bogenschütz, P. Simulation, modeling, and dynamically based parameterization of organized tropical convection for global climate models. *J. Atmos. Sci.* **74**, 1363–1380 (2017).
- Tan, J., Jakob, C., Rossow, W. B. & Tselioudis, G. Increases in tropical rainfall driven by changes in frequency of organized deep convection. *Nature* **519**, 451–454 (2015).
- Stevenson, S. N. & Schumacher, R. S. A 10-year survey of extreme rainfall events in the central and eastern United States using gridded multisensor precipitation analyses. *Monthly Weather Rev.* **142**, 3147–3162 (2014).
- Feng, Z. et al. Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *J. Adv. Modeling Earth Syst.* **10**, 1470–1494 (2018).
- Moeng, C.-H. & LeMone, M. A. Atmospheric planetary boundary-layer research in the US: 1991–1994. *Rev. Geophys.* **33**, 923–931 (1995).
- Smith, J. A., Baeck, M. L., Zhang, Y. & Doswell, C. A. III Extreme rainfall and flooding from supercell thunderstorms. *J. Hydrometeorol.* **2**, 469–489 (2001).
- Cintineo, R. M. & Stensrud, D. J. On the predictability of supercell thunderstorm evolution. *J. Atmos. Sci.* **70**, 1993–2011 (2013).
- Bretherton, C. S., Blossey, P. N. & Khairoutdinov, M. An energy-balance analysis of deep convective self-aggregation above uniform SST. *J. Atmos. Sci.* **62**, 4273–4292 (2005).
- Khairoutdinov, M. F. & Emanuel, K. A., Aggregated convection and the regulation of tropical climate. In *29th Conference on Hurricanes and Tropical Meteorology* (American Meteorological Society, Tucson, AZ, 2010).
- Muller, C. J. & Held, I. M. Detailed investigation of the self-aggregation of convection in cloud-resolving simulations. *J. Atmos. Sci.* **69**, 2551–2565 (2012).
- Held, I. M., Hemler, R. S. & Ramaswamy, V. Radiative-convective equilibrium with explicit two-dimensional moist convection. *J. Atmos. Sci.* **50**, 3909–3927 (1993).
- Tompkins, A. M. & Craig, G. C. Radiative-convective equilibrium in a three-dimensional cloud-ensemble model. *Q. J. R. Meteorological Soc.* **124**, 2073–2097 (1998).
- Hohenegger, C. & Stevens, B. Coupled radiative convective equilibrium simulations with explicit and parameterized convection. *J. Adv. Modeling Earth Syst.* **8**, 1468–1482 (2016).

14. Ruppert, J. H. Jr. & Hohenegger, C. Diurnal circulation adjustment and organized deep convection. *J. Clim.* **31**, 4899–4916 (2018).
15. Ruppert, J. H. Jr. & O'Neill, M. E. Diurnal cloud and circulation changes in simulated tropical cyclones. *Geophys. Res. Lett.* **46**, 502–511 (2019).
16. Jeevanjee, N. & Romps, D. M. Convective self-aggregation, cold pools, and domain size. *Geophys. Res. Lett.* **40**, 994–998 (2013).
17. Haerter, J. O. Convective self-aggregation as a cold pool-driven critical phenomenon. *Geophys. Res. Lett.* **46**, 4017–4028 (2019).
18. Weller, R. A. & Anderson, S. P. Surface meteorology and air-sea fluxes in the western equatorial pacific warm pool during the toga coupled ocean-atmosphere response experiment. *J. Clim.* **9**, 1959–1990 (1996).
19. Johnson, R. H., Rickenbach, T. M., Rutledge, S. A., Ciesielski, P. E. & Schubert, W. H. Trimodal characteristics of tropical convection. *J. Clim.* **12**, 2397–2418 (1999).
20. Kawai, Y. & Wada, A. Diurnal sea surface temperature variation and its impact on the atmosphere and ocean: a review. *J. Oceanogr.* **63**, 721–744 (2007).
21. Tian, B., Waliser, D. E. & Fetzer, E. J. Modulation of the diurnal cycle of tropical deep convective clouds by the MJO. *Geophys. Res. Lett.* **33**, L20704 (2006).
22. Suzuki, T. Diurnal cycle of deep convection in super clusters embedded in the Madden-Julian Oscillation. *J. Geophys. Res.: Atmos.* **114**(D22), 1–14 (2009).
23. Liu, C. & Moncrieff, M. W. A numerical study of the diurnal cycle of tropical oceanic convection. *J. Atmos. Sci.* **55**, 2329–2344 (1998).
24. Chen, S. S., Houze, R. A. Jr. & Mapes, B. E. Multiscale variability of deep convection in relation to large-scale circulation in TOGA COARE. *J. Atmos. Sci.* **53**, 1380–1409 (1996).
25. Droegeleier, K. K. & Wilhelmson, R. B. Three-dimensional numerical modeling of convection produced by interacting thunderstorm outflows. Part I: control simulation and low-level moisture variations. *J. Atmos. Sci.* **42**, 2381–2403 (1985).
26. Rotunno, R., Klemp, J. B. & Weisman, M. L. A theory for strong, long-lived squall lines. *J. Atmos. Sci.* **45**, 463–485 (1988).
27. Tompkins, A. M. Organization of tropical convection in low vertical wind shears: the role of cold pools. *J. Atmos. Sci.* **58**, 1650–1672 (2001).
28. Böing, S. J., Jonker, H. J. J., Siebesma, A. P. & Grabowski, W. W. Influence of the subcloud layer on the development of a deep convective ensemble. *J. Atmos. Sci.* **69**, 2682–2698 (2012).
29. Schlemmer, L. & Hohenegger, C. Modifications of the atmospheric moisture field as a result of cold-pool dynamics. *Q. J. R. Meteorological Soc.* **142**, 30–42 (2016).
30. Feng, Z. et al. Mechanisms of convective cloud organization by cold pools over tropical warm ocean during the amie/dynamo field campaign. *J. Adv. Modeling Earth Syst.* **7**, 357–381 (2015).
31. Moseley, C., Hohenegger, C., Berg, P. & Haerter, J. O. Intensification of convective extremes driven by cloud-cloud interaction. *Nat. Geosci.* **9**, 748 (2016).
32. Haerter, J. O., Böing, S. J., Henneberg, O. & Nissen, S. B. Circling in on convective organization. *Geophys. Res. Lett.* **46**, 7024–7034 (2019).
33. Lochbihler, K., Lenderink, G. & Siebesma, A. P. Response of extreme precipitating cell structures to atmospheric warming. *J. Geophys. Res.: Atmos.* **124**, 6904–6918 (2019).
34. Petch, J. C., Brown, A. R. & Gray, M. E. B. The impact of horizontal resolution on the simulations of convective development over land. *Q. J. R. Meteorological Soc.* **128**, 2031–2044 (2002).
35. Guichard, F. et al. Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Q. J. R. Meteorological Soc.* **130**, 3139–3172 (2004).
36. Sato, T., Miura, H., Satoh, M., Takayabu, Y. N. & Wang, Y. Diurnal cycle of precipitation in the tropics simulated in a global cloud-resolving model. *J. Clim.* **22**, 4809–4826 (2009).
37. Schlemmer, L. *The Diurnal Cycle of Midlatitude, Summertime Moist Convection Over Land in an Idealized Cloud-resolving Model*. PhD thesis, ETH Zurich (2011).
38. Haerter, J. O. & Schlemmer, L. Intensified cold pool dynamics under stronger surface heating. *Geophys. Res. Lett.* **45**, 6299–6310 (2018).
39. Chen, S. S. & Houze, R. A. Diurnal variation and life-cycle of deep convective systems over the tropical pacific warm pool. *Q. J. R. Meteorological Soc.* **123**, 357–388 (1997).
40. Ban, N., Schmidli, J. & Schär, C. Heavy precipitation in a changing climate: does short-term summer precipitation increase faster? *Geophys. Res. Lett.* **42**, 1165–1172 (2015).
41. Prein, A. F. et al. Simulating North American mesoscale convective systems with a convection-permitting climate model. *Clim. Dyn.* **55**, 1–16 (2017).
42. Rasp, S., Selz, T. & Craig, G. C. Variability and clustering of midlatitude summertime convection: testing the craig and cohen theory in a convection-permitting ensemble with stochastic boundary layer perturbations. *J. Atmos. Sci.* **75**, 691–706 (2018).
43. Satoh, M. et al. Global cloud-resolving models. *Curr. Clim. Change Rep.* **5**, 1–13 (2019).
44. Houze, R. A. Jr. Mesoscale convective systems. *Rev. Geophys.* **42**, 1–43 (2004).
45. Yang, G.-Y. & Slingo, J. The diurnal cycle in the tropics. *Monthly Weather Rev.* **129**, 784–801 (2001).
46. Janowiak, J. E., Arkin, P. A. & Morrissey, M. An examination of the diurnal cycle in oceanic tropical rainfall using satellite and in situ data. *Monthly weather Rev.* **122**, 2296–2311 (1994).
47. Held, I. M. & Soden, B. J. Robust responses of the hydrological cycle to global warming. *J. Clim.* **19**, 5686–5699 (2006).
48. Feller, W. *An Introduction to Probability Theory and Its Applications* (John Wiley & Sons, 1957).
49. Böing, S. J. An object-based model for convective cold pool dynamics. *Math. Clim. Weather Forecast.* **2**, 43–60 (2016).
50. Etling, D. *Theoretische Meteorologie: Eine Einführung* (Springer-Verlag, 2008).
51. Zhang, Y. & Klein, S. A. Mechanisms affecting the transition from shallow to deep convection over land: inferences from observations of the diurnal cycle collected at the ARM southern great plains site. *J. Atmos. Sci.* **67**, 2943–2959 (2010).
52. Seifert, A. & Beheng, K. D. A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteorol. Atmos. Phys.* **92**, 45–66 (2006).
53. Christensen, K. & Moloney, N. R. *Complexity and Criticality*, Vol. 1 (World Scientific Publishing Company, 2005).
54. Lenderink, G. & Van Meijgaard, E. Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nat. Geosci.* **1**, 511–514 (2008).
55. Berg, P., Moseley, C. & Haerter, J. O. Strong increase in convective precipitation in response to higher temperatures. *Nat. Geosci.* **6**, 181–185 (2013).
56. Coumou, D. & Rahmstorf, S. A decade of weather extremes. *Nat. Clim. Change* **2**, 491 (2012).
57. Muller, C. & Bony, S. What favors convective aggregation and why? *Geophys. Res. Lett.* **42**, 5626–5634 (2015).
58. Hu, H., Leung, L. R. & Feng, Z. Observed warm-season characteristics of mcs and non-mcs rainfall and their recent changes in the central united states. *Geophys. Res. Lett.* **47**, e2019GL086783 (2020).
59. Smagorinsky, J. General circulation experiments with the primitive equations: I. the basic experiment. *Monthly Weather Rev.* **91**, 99–164 (1963).
60. Pincus, R. & Stevens, B. Monte Carlo spectral integration: a consistent approximation for radiative transfer in large eddy simulations. *J. Adv. Modeling Earth Syst.* **1**, 1–9 (2009).
61. Stevens, B. et al. Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Monthly Weather Rev.* **133**, 1443–1462 (2005).
62. Stull, R. B. *An Introduction to Boundary Layer Meteorology*, Vol. 13 (Springer Science & Business Media, 2012).
63. Moseley, C., Henneberg, O. & Haerter, J. O. A statistical model for isolated convective precipitation events. *J. Adv. Modeling Earth Syst.* **11**, 360–375 (2019).
64. Bretherton, C. S. & Smolarkiewicz, P. K. Gravity waves, compensating subsidence and detrainment around cumulus clouds. *J. Atmos. Sci.* **46**, 740–759 (1989).

## ACKNOWLEDGEMENTS

We thank Kim Sneppen for fruitful discussions on the simplified modelling. J.O.H. and B.M. gratefully acknowledge funding by a grant from the VILLUM Foundation (grant number: 13168) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant number: 771859). S.B. N. acknowledges funding through the Danish National Research Foundation (grant number: DNRF116). The authors acknowledge the Danish Climate Computing Center (DC3) and the German Climate Computing Center (DKRZ).

## AUTHOR CONTRIBUTIONS

J.O.H. ran and processed the large-eddy simulations (LES), developed the conceptual models and wrote the paper. B.M. produced the case study and revised the paper. S. B.N. contributed to model development and implementation and revised the paper.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41612-020-00132-z>.

**Correspondence** and requests for materials should be addressed to J.O.H.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020